Going Beyond Functionings to Capabilities : an Econometric Model to Explain and Estimate Capabilities

Jaya Krishnakumar

No 2004.12

Cahiers du département d'économétrie
Faculté des sciences économiques et sociales
Université de Genève

Août 2004

Département d'économétrie
Université de Genève, 40 Boulevard du Pont-d'Arve, CH -1211 Genève 4
http://www.unige.ch/ses/metri/

Going beyond functionings to capabilities: an econometric model to explain and estimate capabilities

by

Jaya Krishnakumar[1]
University of Geneva

August 2004

Abstract

Any attempt to operationalise the capability approach necessitates an adequate framework for the measurement of the abstract unobservable multidimensional concept that the term human development stands for. One such attempt is the latent variable approach including principal components, factor analysis and MIMIC models. The first two models provide estimates of the latent variables but are silent on the factors influencing these variables (capabilities in our context). MIMIC models represent a step further in this direction as they include exogenous "causal" variables for the latent factors but the effects go only in one direction i.e. from the "causes" to the latent variables. We argue that some of these causal factors not only influence human development but they are also influenced by it and that unless this feedback mechanism is taken into account we do not have a complete picture of this complex phenomenon. In this paper we present a theoretical framework incorporating the above aspects into a coherent system of causes, effects and interactions, leading to an econometric model which represents a generalisation of existing latent variable models. Estimating the model will enable us to explain the level of capabilities, say how they can be best improved, test our theoretical hypotheses and derive estimators that reflect the actual capabilities rather than just the functionings.

Keywords: human development, capability approach, latent variables, qualitative response, simultaneous equations.

JEL Classification codes : C3, I31, O10

[1]Department of Econometrics, University of Geneva, UNI-MAIL, 40, Bd. du Pont d'Arve, CH-1211, GENEVA 4, Switzerland. Email: jaya.krishnakumar@metri.unige.ch.

Going beyond functionings to capabilities: an econometric model to explain and estimate capabilities

Abstract

Any attempt to operationalise the capability approach necessitates an adequate framework for the measurement of the abstract unobservable multidimensional concept that the term human development stands for. One such attempt is the latent variable approach including principal components, factor analysis and MIMIC models. The first two models provide estimates of the latent variables but are silent on the factors influencing these variables (capabilities in our context). MIMIC models represent a step further in this direction as they include exogenous "causal" variables for the latent factors but the effects go only in one direction i.e. from the "causes" to the latent variables. We argue that some of these causal factors not only influence human development but they are also influenced by it and that unless this feedback mechanism is taken into account we do not have a complete picture of this complex phenomenon. In this paper we present a theoretical framework incorporating the above aspects into a coherent system of causes, effects and interactions, leading to an econometric model which represents a generalisation of existing latent variable models. Estimating the model will enable us to explain the level of capabilities, say how they can be best improved, test our theoretical hypotheses and derive estimators that reflect the actual capabilities rather than just the functionings.

1 Introduction

According to Nobel Prize Laureate Amartya Sen, the basic purpose of development is to enlarge people's choices so that they can lead the life they want to (Sen (1985, 1999)). He also emphasizes that development is a multidimensional concept enveloping diverse social, economic, cultural and political dimensions and that economic growth, though necessary, is not sufficient in itself to bring about development in this large sense.

In Sen's approach, the *choices* that one has are termed "capabilities" and the actual levels of achievement attained in the various dimensions are called "functionings". Thus human development is given by (the enhancement of) the set of choices or capabilities whereas functionings are a set of "beings" and "doings" for example the level of education, the state of health and the extent of participation in the political process. The concept of human development proposed by Mahbub ul Haq, in the first Human Development Report in 1990 (see UNDP(1990)), largely inspired from Sen's various works (e.g. Sen (1985,1999)), represents a major step ahead in the concretization of this extended meaning of development and in the effort to bring people's lives into the center of thinking and analysis. Since then, human development has been the object of extensive theoretical and empirical researches. It has been studied from various angles, conceptual, methodological, operational and policy-making. One such aspect is the measurement issue which is

crucial for a comparative assessment of different situations. As it is not possible to directly observe and measure human development in its large sense, it is generally constructed as a composite index based on several variables (indicators).

The most well-known of these are the Physical Quality of Life Index (PQLI) proposed by Morris (1979) and the Human Development Index (HDI) proposed by the UNDP (1990). The former takes into account life expectancy at age one, infant mortality and adult literacy and the latter the following: life expectancy at birth, an education index (a composite index combining adult literacy rate and school enrolment ratio with a weight of 2/3 and 1/3 respectively) and the real GDP par capita. These indices are given equal weights in the construction of the HDI. Over the recent years other indices came to be proposed which are derived from an underlying theoretical model, that offer some explanation of the variables composing the index as well as a better justification for the choice and values of the weights in the construction of the index.

Income or consumption still remains the most widely used indicator of well-being but it is also one of the most criticised for not capturing the non-economic dimensions of human life (without denying the importance of the economic aspect, cf. Noorbakhsh (1998) and Osberg and Sharpe (2003)). There are ample examples to show that economic growth though necessary

is not sufficient to achieve a good quality of life[2] in various spheres such as the political one (for instance regarding the capability to express one's opinion freely), in the area of personal safety/security (being able to move about freely without being assaulted/arrested, having the right to a fair trial) and many others. In Sen's capability approach (Sen (1999)), the freedom that one enjoys in being able to choose the life one wants is multi-dimensional in nature and economic welfare is only one of the many dimensions it comprises.

A theoretical framework that is appealing in this context is a model which assumes that the capabilities are unobservable latent variables observed through a set of indicators. Principal components, factor analysis and MIMIC (multiple indicators and multiple causes) models all fall into this line of reasoning (cf. Nagar and Basu (2001), Lelli (2001), Biswas and Caliendo (2002), Rahman *et al.* (2003), McGillivray (2003)). Latent variable models are common in psychology and the reader can find an excellent coverage of most of these models with applications in Bartholomew and Knott (1999) and Skrondal and Rabe-Hesketh (2004). The principal components estimate the latent variables as linear combinations of the observed indicators chosen in such a way as to reproduce the original data as closely as possible. But this method lacks an underlying theoretical model which the factor analysis offers. In the latter model the observed values are postulated to be (linear) functions of a certain number (fewer) of unobserved latent variables (called

[2]Throughout this paper we will use the terms 'human development', 'well-being' and 'quality of life' in an interchangeable manner.

factors). Thus it provides a framework for going beyond functionings to reach the capabilities represented by the latent factors. However this model does not *explain* the latent variables (or the capabilities in our context) in that it does not say what causes these capabilities to change. We believe it is as important to be able to say something about the capabilities as it is to say how we can enhance them and thus promote human development. It is not enough to be able to measure how much is achieved but it is also important to be able to say how things can be improved.

The MIMIC model (cf. Joreskog and Goldberger (1975)) represents a step further in the explanation of the phenomenon under investigation as it is not only believed that the observed variables are manifestations of an underlying unobserved latent concept but also that there are other exogenous variables that "cause" and influence the latent factor(s). This structure is highly relevant in our context as there are several institutional, political and social arrangement factors which definitely influence human development and need to be taken into account. Not only do these factors influence human development but they are also influenced by it. A simple example is that if access to education is facilitated, i.e. knowledge capability is increased, development improves and this may in turn incite people to demand free access to education for all (at least in a democratic setting) forcing the government to implement such a policy. This is because there is some sort of a virtuous cycle that is generated by the process of development. Adequate

institutional setups can promote development but it is also true that development in turn encourages favourable political and social arrangements by making people more and more aware, involved and demanding and enforces the participatory element of progress. Thus there is a feedback mechanism by which human development promotes its own "causal" factors. Unless this feedback mechanism is taken into account we do not have a complete picture of the evolving nature of the whole system.

There are other models in the psychometric literature such as LISREL with ordinal variables (cf. Muthen (1983, 1984) and Arminger and Küsters (1988), Joreskog (2002)) and MIMIC with exogenous variables (cf. Moustaki (2003)) that represent useful extensions for our context (though not yet applied in this field to our knowledge). However, as we argue in the following sections none of them seem to incorporate all the features that we believe are essential for adequate modelling of the capability approach. For instance MIMIC has exogenous causes of latent variables but lacks structural interdependence; LISREL and its generalisations account for simultaneity but lack exogenous elements in measurement modules.

In this paper we propose a theoretical framework that takes all the relevant aspects into account in an appropriate way. Then we transform it into an econometric model which can be fitted using real data enabling a better understanding of how this complex mechanism works in practice. It would

also allow us to verify our assumptions about the feedback mechanism mentioned above. Finally it would give us estimates of the actual capabilities rather than just the functionings.

The next section puts forward the case for the interdependent nature of capabilities by taking some of the most important components of human welfare such as education, health and social participation. In Section 3 we bring in the measurement relations based on our postulate that capabilities are latent and manifest themselves in the form of functionings. Arguments of Sections 2 and 3 combined will provide us with the necessary foundation for formulating our theoretical framework in Section 4. This will in turn lead to the econometric model presented in Section 5 where we also briefly touch upon estimation issues. Section 6 ends the paper with some concluding remarks.

2 The Simultaneous Nature of Capabilities

We mentioned earlier that 'capabilities' are the choices that one faces in life and 'functionings' are the outcomes. Then it is not difficult to imagine that there could be more than one achievement level for the same capability level. Take education for instance. The 'capability' in this field is given by the freedom to increase one's knowledge through education which is in turn facilitated by access to a good school. However one person may exercise the

freedom by actually going to school and getting educated whereas another may use the same choice in not going to school due to various reasons. Thus we need a framework in which the same level of capability can give rise to different outcomes depending on external factors (individual, social and environmental) influencing the decision-making process. Formally, this would mean that some exogenous variables need to be added in the system of equations linking the observed response (functioning) to the latent capability.

Let us go further with the same example to get an idea of what these exogenous factors could be. Considering the education of a child in a developing country (especially in rural areas), family perceptions of the return on education compared to the immediate consequence of helping at home or in the field could play a role in deciding whether to send the child to school or not, independent of the availability of a school in the village. Though there is the subtle point that the child may not have the choice here, it is beyond the scope of the present paper to go deeper into this issue. Here we take the view that there is a choice but it is restricted by family compulsions. Another crucial element which comes into play in most developing countries is the gender of the child. Unfortunately it is still not uncommon that only boys are given proper education in certain traditions and girls are excluded from the process as boys are seen as income-earners who stay with the parents for ever thus adding to the total household income and ensuring that parents are taken care of in their old age. On the other hand, the family can also give

importance to the non-monetary benefits of education (of its children) which will lift its status in the society as learned persons always command more respect (wealth is no doubt another important contributor to the social status and education helps here too by providing better job opportunities). Needless to mention the value added to one's personality and the self-confidence procured by knowledge acquired through education. Thus we see that several personal or subjective characteristics enter the decision-making process sometimes acting in opposing directions.

Next, let us take health. None can deny the significance of good health as an important constituent of one's well-being. Being healthy is not only an integral part of welfare but also acts as an instrument in enhancing one's capacity to work and earn a living. However all individuals may not react the same way when faced with a health issue. Even assuming that adequate means and infrastructure exist and are accessible, people may choose different options depending on circumstances. Some may go to a public health centre, some to a private one. Some may not avail of these professional services but instead may follow a more traditional route of consulting a family/social guide in this matter, a custom still prevalent in many rural areas. In such situations, there is bound to be a difference in the result given the same choice depending on one's own convictions, social traditions, family practices and on the degree of acceptance of alternative forms of medicine which are more and more sought after in developed countries too.

Taking a different angle, one can argue that education brings about a better awareness of health and environmental issues and enables one to think of options that may not have even been part of the choice set otherwise. This is actually equivalent to saying that it increases the range of choice i.e the capability set itself. For instance it is well-known that educating a mother has a direct impact on her own and her children's health and well-being (cf. Murthi, Guio and Drèze (1997) e.g.) meaning that there is a clear interaction between education and health. Thus improving one capability can affect another in a favourable manner implying that capabilities are interdependent and this property should be included in the theoretical model we are trying to develop.

Let us remain in the health domain and consider yet another aspect. We mentioned earlier that health is valuable not only in itself but also in enabling a person to be usefully occupied (whether it be for earning a livelihood for oneself and one's family or for helping others). Also, the healthier one is, the more active one can be in participating in local community affairs on the political, social or environmental fronts. Once again this will positively influence the choice-set on the whole. Imagine a poor area in a developing country where there is lack of safe drinking water supply. An active involvement of local citizens is sometimes the only way to alert the otherwise indifferent and/or corrupt political authorities. For this to happen, the citizens must have the necessary knowledge, exposure and health to be able to organise

themselves and exercise their political and social rights. One can of course argue that certain institutions need to be in place for action to be pursued in this regard. Well, if they do not exist then the local citizens may even end up taking the initiative to create them. This only goes to show that many 'exogenous' factors affecting capabilities (the institutional setup in this example) can in turn be affected by them i.e they are not exogenous at all! This is the feedback mechanism that we mentioned in the introduction and one cannot ignore this simultaneous nature of our variables. At the same time it is also true that there are some purely exogenous factors like the traditions that we talked about in our previous paragraphs on education and health or in the water supply example it could be the existence of a river or a lake nearby and/or rain water storage facilities etc.

One can go on and on with many other arguments to support the case for the interdependent nature of capabilities but we believe there is no further need to elaborate on this. Not only do capabilities interact among themselves but also with other elements representing the socio-political setup. For some elements belonging to latter group, there are feedback effects (thus making them jointly dependent) whereas for others the causal link only operates in one direction (making them purely exogenous).

3 The Measurement Issue

Capabilities by definition cannot be directly measured. Hence they need to be specified as latent unobservable variables in our model. What can be measured however are the functionings namely the achievements in each dimension both at the individual (household) and at the national levels. These achievements are generally identified by proper indicators reflecting the performance in the associated dimension. There could either be one indicator or as is more often the case a whole range of indicators available for each capability dimension. In other words, one normally has a vector of functionings rather than a scalar indicator corresponding to each domain. In the case of health, at the national level, one can think of indicators such as life expectancy, infant/child mortality, total fertility, number of doctors for 1000 persons, number of hospital beds for 1000 persons and so on. Sometimes one may need to combine all of these to give a single measure. However combining raises additional issues like common units, weights etc. and we do not intend to go deeper into these problems in this paper. At this stage we will limit ourselves to admitting our preference for using vectors so that these problems can be avoided. However we do not exclude the possibility of being forced to aggregate during the practical implementation or estimation of our model due to size limitations that come up while running the program. Examination of this issue is therefore left for the future.

There are several types of indicators available in practice. Some of them could be continuous like the above-mentioned life expectancy, per capita number of doctors whereas some could be of a qualitative nature for instance the existence of the right to vote or not, existence of safe water access or not, a school or a hospital in the neighbourhood or not, existence of adequate sanitation facilities or not. At the individual level one could also have *subjective* assessments such as whether a person considers herself to be poor or not. The above characteristics are examples of what is called a binary or dichotomous variable (with two possible outcomes: yes and no coded as either 1 and 0 or +1 and -1). There are also other types of qualitative indicators: polychotomous (more than two outcomes e.g. different levels of education - no formal education, primary, secondary, college...). Note that there is a certain order in the last variable and hence it is termed as an ordinal variable. There could also be polychotomous variables with no order for example religion - Hindu, Muslim, Buddhist, Christian etc. Some other indicators could be truncated or censored - truncated when not observed for a particular range of values, censored when observed only if greater than a threshold value. One should bear in mind that the statistical/econometric treatment of these variables differs according to the particular type concerned.

Having established the interdependent nature of the underlying latent capabilities and the observable nature of the outcomes or functionings, it is fundamental that we maintain both sets of variables in our model and link

the two through a set of relationships. In the psychometric literature, these relationships are called 'measurement equations' and the observed outcomes 'response variables'. These will complete our theoretical setting while paying heed to our concern for differentiating between capabilities and functionings.

4 The General Theoretical Framework

Let us recall from the foregoing discussion that the following features need to be present in our framework which should above all *explain* the capabilities:

(i) Capabilities are *latent, unobservable* and interdependent, and are *endogenous* in our structural model.

(ii) Capabilities are influenced by a set of social, political and institutional factors some of which may in turn be influenced by them. (In addition to capabilities there are also some *observed endogenous* variables in our model.)

(iii) Capabilities are also influenced by a set of observable external/*exogenous* causes (such as traditions, cultural elements, natural environmental factors and some social, political, institutional ones which are not part of (ii)).

(iv) Achievements/functionings are measurable and are linked to the underlying capabilities (the set of relationships linking the two is the so-called

measurement model or the qualitative response model).

(v) The relationships between the latent capabilities and the observed functionings are also affected by *exogenous* elements (for instance individual characteristics).

Let us now introduce some notations which will help us formulate our theoretical framework in precise terms. We shall denote by

y^* a vector of latent unobserved capabilities say $(m \times 1)$

y a vector of observed indicators representing the functionings associated with the capability vector say $(p \times 1)$;
as discussed earlier, some these y's could be continuous, some qualitative or discrete

z a vector of observed variables that influence the capabilities but are also influenced by them say $(n \times 1)$

x a vector of exogenous "causes" of y^* and z say $(k \times 1)$

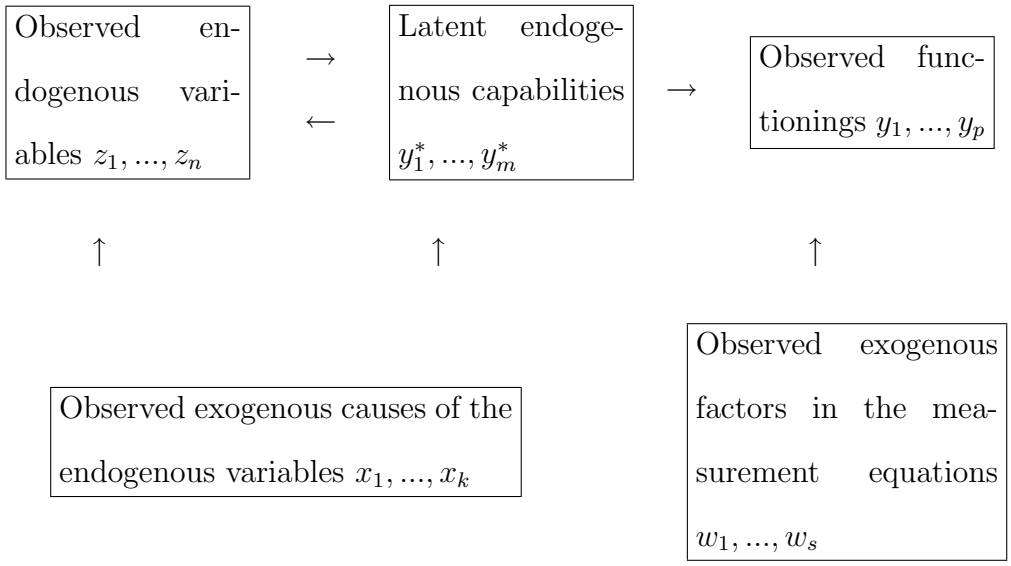
w a vector of exogenous factors entering the measurement equations i.e. the relationships between observed indicators y and latent variables y^* say $(s \times 1)$

For each vector, a typical element will be denoted using a subscript i, e.g. y_i^*, $i = 1, ..., m$.

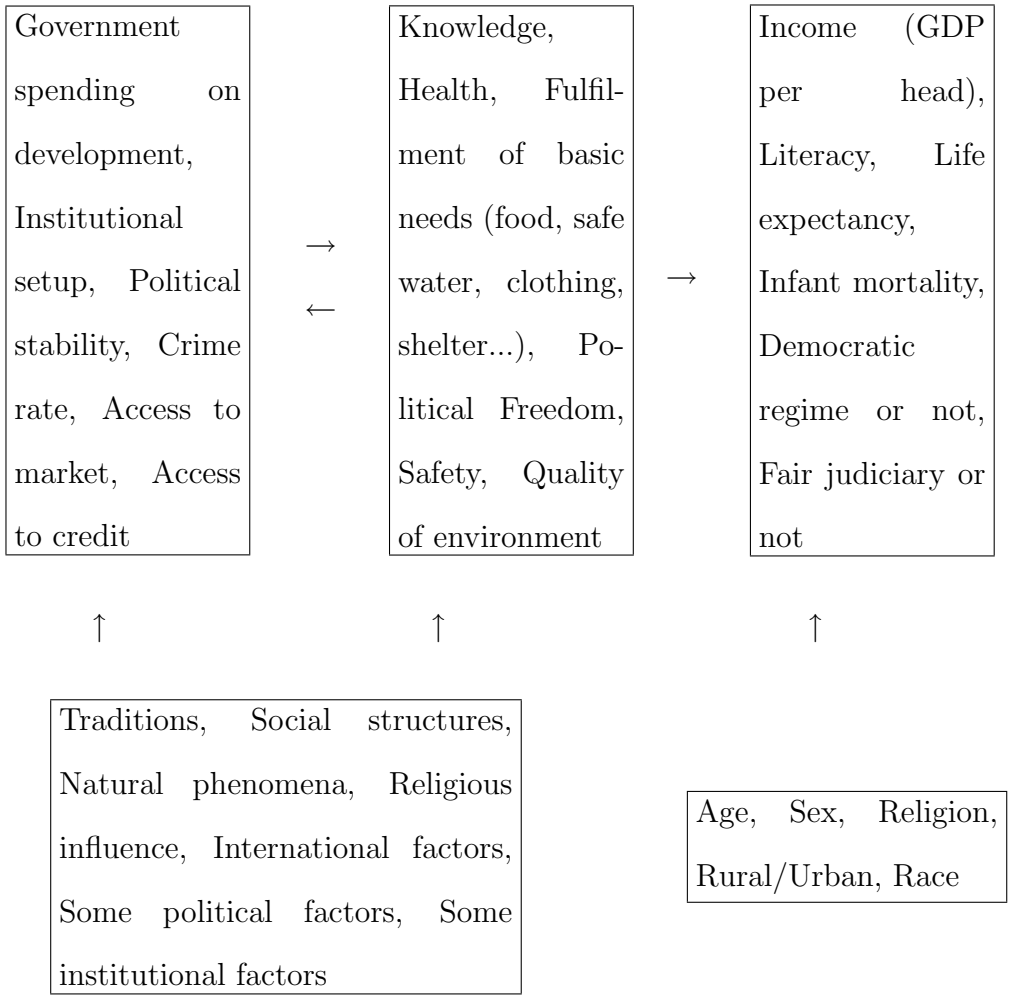
Note that we do not have latent exogenous variables though theoretically it is perfectly possible to allow for such a case. The reason for not including

them in the above framework is that we do not see their relevance in our practical context where we would normally directly observe all exogenous factors.

Keeping all the above features in mind we can represent our structure by the following diagram (which is usually called the path diagram in the social science literature).



For the reader to get a concrete idea of how this framework operates in practice we will reproduce the above diagram below replacing the abstract variables in each box by some examples.



5 The Econometric Model

The conceptual framework arrived at in Section 4 leads us to a general mixed (latent and observed) simultaneous equation model which we can now write down in formal terms as follows.

$$Ay^* + Bz + Cx + u = 0 \tag{1}$$

$$g(y) = h(y^*, w) + v \tag{2}$$

The first set of equations represents the structural model which jointly explains (y^*, z) in terms of x, with A, B, C being the corresponding coefficient matrices of appropriate dimensions. We have used the term 'mixed' to indicate that there are both *latent endogenous* y^* (with qualitative response for some) and *observed endogenous* z (continuous) variables in our structural simultaneous equation model (SEM)[3]. The second set of equations forms the measurement model or the qualitative response model (QRM) where it is specified how the latent variables are related to the observed responses through functions $g(\cdot)$ and $h(\cdot)$. Note the presence of exogenous variables in both the models.

Vectors u and v are the respective error vectors in the SEM and QRM, with zero expectations, uncorrelated between the two parts but correlated within each. Let us denote

$$V(u) = \Sigma$$

and

$$V(v) = \Psi.$$

[3]Some authors (Bartholomew and Knott (1999), Moustaki (2003)) use the same term to denote a mixture of different types of qualitative responses.

In general Ψ is assumed to be diagonal in the latent variable model literature. Further, depending on the nature of y variance of some elements of v will be specified as unity (for proper identification of the coefficients).

As far as the SEM part is concerned, certain elements of the coefficient matrices (those appearing in the structural equations explaining the latent variables) can only be estimated up to a proportionality factor under the usual identification conditions. The reader is referred to Maddala (1983) for further explanations.

It is interesting to observe that this general model includes many known models as "special cases" which we will come to shortly. Before that, we also note that there are some other cases where both y^* and y appear in the SEM part that fall outside this framework (cf. Maddala (1983)). In such cases one has to pay attention to the additional problem of logical consistency or coherence of the whole system. This problem does not arise in our model in so far as the SEM part only contains the latent variables without their observed counterparts. Though we can also theoretically extend our framework to such situations, we are not sure of their practical relevance in our context where it would mean that capabilities lead to functionings but functionings influence capabilities too. If we argue that functionings are measurable achievements given the level of capabilities, there is no sense in assuming that they affect capabilities (they should simply reflect how 'well'

the capabilities are converted to actual achievements). In our opinion, it is more meaningful to make capabilities mutually dependent and keep the relationship between capabilities and functionings just one way.

Let us now identify some special cases of our model that are of interest in our field of application.

Case 1

If y is continuous, $g(\cdot), h(\cdot)$ linear and there is no w we get the standard LISREL model (cf. Joreskog (1973)) (with observed rather than latent exogenous, refer to an earlier remark in this respect).

Case 2

With ordinal y and no w we have LISREL with ordinal variables (cf. Joreskog (2002), Muthen (1983, 1984)). The latter author has two types of measurement equations: 'inner' measurement equations and 'outer' measurement equations as he allows for *latent* response variables and *observed* response variables.

Case 3

If y^* scalar, $A = 1$, no z, no w, y continuous we have the MIMIC model (cf. Joreskog and Goldberger (1975)).

Case 4

Same as Case 3 with y^* a vector, $A = I$, we have the extended or generalised MIMIC.

Case 5

Same as Case 4 with w and z, we have the MIMIC with covariates (cf. Moustaki (2003)).

Case 6

If y^* is observed (no measurement equation) then we have the classical SEM (cf. e.g. Theil (1979), Hausman(1983)).

Case 7

If y^* is observed and $A = I$, then we have the SUR model (cf. Zellner (1962)).

Case 8

When y^* is scalar (no z) and y is either discrete or limited dependent we have the classical qualitative dependent variable model (see Amemiya (1985)).

In absence of any of these special cases, we have the general mixed simultaneous equation model as mentioned earlier. Though these types of models

are discussed in Maddala (1983) he does not include the possibility of some latent variables having continuous responses. His model has some latent and some observed endogenous variables like ours but the latent variables are all observed as qualitative responses. In general these models can be estimated by two-stage methods as described in Maddala (1983). First a reduced form ML (probit if dichotomous, tobit if censored etc.) - univariate or bivariate or multivariate - is performed to get estimates of y^* to be used in the structural form which can then be estimated by ML for the latent ones and IV/GMM methods for the observed z[4]. All these estimates are consistent and asymptotically normal though the asymptotic standard errors obtained in the second stage have to be corrected for the heteroscedasticity resulting from the fact some of the explanatory variables are estimated. We will not go deeper into the estimation issues in this paper. They will be discussed in detail in the application paper to follow.

Let us illustrate how the above model functions in a particular case with two dimensions namely education and health. Here we would like to stress that this example is only for illustration purposes and should not be taken as resulting from a serious study of the subject. We have two latent capabilities represented by $y^* = [y_1^* \, y_2^*]' = [\text{knowledge health}]'$. Now what are the indicators available? Let us suppose we have individual data on the level of education (educ), number of visits to the doctor during a certain interval of

[4]In fact one might employ methods similar to the ones suggested by Muthen (1983,1984) for LISREL with categorical data for our reduced form.

time (doc) and the body mass index (bmi)[5]. The first one is the achievement in the education field and the remaining two are those in the health field. So we have $y = [y_1\, y_2\, y_3]' = [\text{educ doc bmi}]'$. Note that y_1 is ordinal and y_2, y_3 are continuous.

In addition suppose we have information on age, sex, number of members in the household, rural/urban, household type, religion, number of males/females, number of children, income (or total consumption), access to safe water, access to school, access to hospital etc. Then we could have $z = [z_1\, z_2]' = [\text{number of children, income}]'$ which is potentially endogenous as the number of children may have a negative influence on the level of education whereas the level of education in turn may also negatively influence the number of children. Similarly a higher income may favour a higher level of education or better health just as it may be the result of higher education or better health.

Finally x could consist of a constant, access to safe water, access to school, access to hospital, household size, household type, religion and w could consist of a constant, age, sex, religion, rural/urban.

Thus our model could be specified as

[5]BMI is defined as mass (in kg) divided by height^2 (in m^2).

$$
\begin{aligned}
y_1^* + A_{12}y_2^* + B_{11}z_1 + B_{12}z_2 + C_{11}x_1 + C_{13}x_3 & \\
+C_{16}x_6 + C_{17}x_7 + u_1 &= 0 \qquad (3)\\
y_2^* + A_{21}y_1^* + B_{22}z_2 + C_{11}x_1 + C_{12}x_2 + C_{14}x_4 & \\
+C_{15}x_5 + C_{26}x_6 + u_2 &= 0 \qquad (4)\\
z_1 + A_{31}y_1^* + B_{32}z_2 + C_{31}x_1 + C_{35}x_5 + C_{37}x_7 + u_3 &= 0 \qquad (5)\\
z_2 + A_{41}y_1^* + A_{42}y_2^* + C_{41}x_1 + C_{45}x_5 + C_{46}x_6 + u_4 &= 0 \qquad (6)
\end{aligned}
$$

$$
\begin{aligned}
p\left(y_1 = j\right) &= p\left(t_j < y_1^* + w'\gamma_1 < t_{j+1}\right) \qquad (7)\\
\begin{bmatrix} y_2 \\ y_3 \end{bmatrix} &= y_2^*\,\alpha + w'\,\gamma_2 + v \qquad (8)
\end{aligned}
$$

where $t_j, j = 1, ..., c$ (number of categories) represent the threshold values and where the reduced form equation is to be substituted for y_1^* in (7).

Note that not all exogenous and endogenous variables appear in all the structural equations and one endogenous variable is normalised in each equation (this not only has economic sense but is also required for identification purposes).

Now, how would one go about estimating such a model? Following the

two stage procedure, one would first estimate the reduced form of the SEM part given by

$$
\begin{aligned}
y_1^* &= \pi_{11}x_1 + \pi_{12}x_2 + \pi_{13}x_3 + \pi_{14}x_4 + \pi_{15}x_5 + \pi_{16}x_6 + \pi_{17}x_7 + \tilde{u}_1 \\
y_2^* &= \pi_{21}x_1 + \pi_{22}x_2 + \pi_{23}x_3 + \pi_{24}x_4 + \pi_{25}x_5 + \pi_{26}x_6 + \pi_{17}x_7 + \tilde{u}_2 \\
z_1 &= \pi_{31}x_1 + \pi_{32}x_2 + \pi_{33}x_3 + \pi_{34}x_4 + \pi_{35}x_5 + \pi_{36}x_6 + \pi_{17}x_7 + \tilde{u}_3 \\
z_2 &= \pi_{41}x_1 + \pi_{42}x_2 + \pi_{43}x_3 + \pi_{44}x_4 + \pi_{45}x_5 + \pi_{46}x_6 + \pi_{17}x_7 + \tilde{u}_4
\end{aligned}
$$

The last two equations of the above system can be estimated by straightforward least squares (or ML) methods whereas the first two require a bit more care. The first one should be estimated by multinomial probit or logit using its measurement equation (7). The second one would be estimated as a MIMIC type model using its measurement equations (8). Once these reduced form equations are estimated one can get estimates of y_1^* and y_2^* and use them as explanatory variables in the structural model (instead of the latent ones) and estimate each equation according to its nature i.e. by multinomial probit for the first structural equation, MIMIC type for the second one, and IV methods for the third and the fourth equations. Finally using the estimates of the structural coefficients one can rederive the reduced form coefficients and use them to predict the latent capabilities.

6 Concluding Remarks

In this paper we have presented a framework that not only distinguishes between capabilities and functionings but also provides an explanation of the level of capabilities both in terms of endogenous and exogenous factors. That different functionings are compatible with the same level of capabilities is made possible by the exogenous variables in the qualitative model. Further, one can also predict capability levels using our model.

Having developed the right setting for analysing and assessing capabilities, the next natural step is the actual implementation of the model with real data. One can then compare our model results with those given by some of the special cases mentioned above such as the extended MIMIC or LISREL with ordinal variables. This is precisely what we propose to do in the future!

References

Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge, Massachusetts.

Arminger G. and U. Küsters (1988), "Latent Trait Models Indicators of Mixed Measurement Level", in R. Langeheine and J. Rost (eds.), *Latent Trait and Latent Class Models* (pp. 51-73), Plenum Press, New York.

Bartholomew D.J. and M. Knott (1999), *Latent Variable Models and*

Factor Analysis, Edward Arnold, U.K.

Biswas B. and F. Caliendo (2002), "A Multivariate Analysis of the Human Development Index", *Indian Economic Journal*, Vol. 49, No. 4.

Hausman, J.A. (1983) "Specification and Estimation of Simultaneous Equation Models", in *Handbook of Econometrics*, North Holland, Amsterdam.

Jöreskog, K. (1973), "A General Method for Estimatimg a Linear Structural Equation system", in A. S. Goldberger and O.D. Duncan (eds), *Structural Equation Models in the Social Sciences*, Seminar Press, New York.

Jöreskog, K. (2002), "Structural Equation Modelling with Ordinal Variables using LISREL", http://www.ssicentral.com/lisrel/ordinal.htm

Jöreskog K. and A. Goldberger (1975), "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable", *Journal of the American Statistical Association*, Vol. 70, No. 351.

Lelli S. (2001), "Factor Analysis V.S Fuzzy Sets Theory : Assessing the influence of different techniques on Sen's Functioning Approach", Center for Economic studies, K.U. Leuven.

Maddala, G.S. (1983), *Limited Dependent and Qualitative variables in Econometrics*, Cambridge University Press, U.K.

McGillivray, M. (2003), "Capturing Non-Economic Dimensions of Human Well-being", Working Paper, WIDER, Helsinki, Finland.

Morris, Morris D. (1979), *Measuring the Condition of the World's Poor: The Physical Quality of Life Index*, Pergamon, New York.

Moustaki, I. (2003) "A General Class of Latent Variable Models for Ordinal Manifest Variables with Covariate Effects on the Manifest and Latent Variables", Dept. of Statistics, Athens University of Economics and Business.

Muthen, B. (1983), "Latent Variable Structural Equation modelling with Categorical Data", *Journal of Econometrics*, 22, 43-65.

Muthen, B. (1984), "A General Structural Equation Model with Dichotomous, Ordered Categorical and Continuous Latent Indicators", *Psychometrika*, 49, 115-132.

Murthi, M, A. Guio, and J. Drèze (1997), "Mortality, fertility, and gender bias in India", in J. Drèze and A. Sen (eds.) *Indian Development: Selected Regional Perspectives*, Oxford University Press, Delhi.

Nagar A.L and S. Basu (2001), "Weighting Socio-Economic Indicators of Human Development (A Latent Variable Approach)", National Institute of Public Finance and Policy, New Delhi.

Noorbakhsh F. (1998), "A Modified Human Development Index", *World Development*, Vol. 26, No. 3, pp. 517-528.

Osberg L. and A. Sharpe (2003), "Human Well Being and Economic Well Being : What Values are Implicit in Current Indices?", WIDER Conference on Inequality, Poverty and Human Well Being, Helsinki 30-31 May 2003.

Rahman, T., R.C. Mittelhammer and P. Wandschneider (2003), "Measuring the Quality of Life across Countries: A Sensitivity Analysis of Well-being Indices", WIDER International Conference on Inequality, Poverty and Human Well-being, Helsinki, Finland.

Sen A.K (1985), *Commodities and Capabilities*, North-Holland, Amsterdam.

Sen A.K. (1999), *Development as Freedom*, Oxford : Oxford University Press.

Skrondal, A. and S. Rabe-Hesketh (2004), *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*, Chapman & Hall/CRC, Boca Raton, U.S.A.

Theil, H. (1979), *Principles of Econometrics*, North Holland, Amsterdam.

UNDP (annual, from 1990), Human Development Report, Oxford University Press, U.K.

Zellner, A. (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", *Journal of the American Statistical Association*, 57, 348-368.

Publications récentes du Département d'économétrie

pouvant être obtenues à l'adresse suivante :

Université de Genève
UNI MAIL
A l'att. de Mme Caroline Schneeberger
Département d'économétrie
40, Bd du Pont-d'Arve
CH - 1211 Genève 4

ou sur

INTERNET : http//www.unige.ch/ses/metri/cahiers

2004.11 MÜLLER Tobias, RAMSES Abul Naga, KOLODZIEJCZYK Christophe, The Redistributive Impact of Altrnative Income Maintenance Schemes : A Microsimulation Study using Swiss Data, Août 2004, 40 pages.

2004.10 GHIGLINO Christian, DUC François, Expectations in an OG Economy, Août 2004, 21 pages.

2004.09 GHIGLINO Christian, OLSZAK-DUQUENNE Marielle, On the Impact of Heterogeneity on Indeterminacy, Août 2004, 29 pages.

2004.08 FIELD Chris, ROBINSON John, RONCHETTI Elvezio, Saddlepoint Approximations for Multivariate M-Estimates with Applications to Bootstrap Accuracy, Août 2004, 26 pages.

2004.07 VICTORIA-FESER Maria-Pia, CONNE David, A Latent Variable Approach for the Construction of Continuous Health Indicators, Août 2004, 8 pages.

2004.06 DELL'AQUILA Rosario, RONCHETTI Elvezio, Resistant Nonparametric Analysis of the Short Term Rate, Juillet 2004, 26 pages.

2004.05 DELL'AQUILA Rosario, RONCHETTI Elvezio, Stock and Bond Return Predictability : The Discrimination Power of Model Selection Criteria, Juin 2004, 25 pages.

2004.04 MANCINI Loriano, RONCHETTI Elvezio, TROJANI Fabio, Optimal Conditionally Unbiased Bounded-Influence Inference in Dynamic Location and Scale Models, Juin 2004 (révision), 35 pages.

2004.03 CANTONI Eva, RONCHETTI Elvezio, A robust approach for skewed and Heavy-tailed outcomes in the analysis of health care expenditures, Mai 2004, 18 pages.

2004.02 MOUSTAKI Irini, VICTORIA-FESER Maria-Pia, Bounded-Bias Robust Estimation in Generalized Linear Latent Variable Models, Avril 2004, 45 pages.